

①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ Patentschrift
⑪ DE 3546807 C2

⑤1 Int. Cl. 5:
C12N 15/10

②1 Aktenzeichen: P 35 46 807.6-41
②2 Anmeldetag: 17. 6. 85
④3 Offenlegungstag: —
④5 Veröffentlichungstag
der Patenterteilung: 28. 3. 91

DE 3546807 C2

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

③0 Unionspriorität: ③2 ③3 ③1
30.03.85 CH 01379/85-8

⑦3 Patentinhaber:
Ballivet, Marc, Genf/Genève, CH; Kauffman, Stuart
Alan, Bryn Mawr, Pa., US

⑦4 Vertreter:
Haft, U., Dipl.-Phys., 8000 München; Berngruber, O.,
Dipl.-Chem. Dr.rer.nat., 8232 Bayerisch Gmain;
Czybulka, U., Dipl.-Phys., Pat.-Anwälte, 8000
München

⑥2 Teil aus: P 35 90 766.5

⑦2 Erfinder:
gleich Patentinhaber

⑤6 Für die Beurteilung der Patentfähigkeit
in Betracht gezogene Druckschriften:
NICHTS ERMITTELT

⑤4 Verfahren zur Herstellung von RNA

DE 3546807 C2

Beschreibung

Die Erfindung betrifft das im Anspruch 1 angegebene Verfahren zur Herstellung von RNA mit einer vorgegebenen Eigenschaft. Die Ansprüche 2 bis 4 betreffen Ausgestaltungen dieses Verfahrens.

Dabei wird nach rekombinanter DNA-Technologie vorgegangen. Bei der bekannten rekombinanten DNA-Technologie wird zur Herstellung des Gens, das in die Wirtszellen eingebracht wird, entweder natürliche DNA zerschnitten, um fremde DNA-Fragmente zu bilden, oder es wird eine bestimmte fremde DNA aus einzelnen Nukleotiden synthetisiert. Die fremde DNA, aus dem das Gen hergestellt wird, das in die Wirtszellen eingebracht wird, weist also in beiden Fällen eine definierte Nukleotid-Sequenz auf.

Aufgabe der Erfindung ist es, ein Verfahren bereitzustellen, mit dem eine RNA erzeugt werden kann, die gegenüber den in der Natur vorkommenden RNAs überlegene Eigenschaften aufweist.

Dies wird erfindungsgemäß mit dem im Anspruch 1 gekennzeichneten Verfahren erreicht. In den Ansprüchen 2 bis 4 sind bevorzugte Ausgestaltungen des erfindungsgemäßen Verfahrens wiedergegeben.

Nach der Erfindung können neue RNA-Sequenzen mit brauchbaren Eigenschaften, insbesondere chemischen, biochemischen oder biologischen Eigenschaften, hergestellt werden.

Es ist deshalb klar, daß die Erfindung eine sehr große Zahl von Anwendungen auf sehr vielen Gebieten der Wissenschaft, Industrie und Medizin eröffnet.

Die stochastischen Gene sind z. B. herstellbar durch stochastische Copolymerisation der vier Arten von Deoxyphosphonukleotiden A, C, G und T, von den zwei Enden eines anfänglich linearisierten Expressionsvektors, gefolgt von der Bildung kohäsiver Enden in einer solchen Weise, um einen stochastischen ersten Strang von DNA zu bilden, der von einem Molekül des Expressionsvektors gebildet wird, das zwei stochastische Sequenzen, dessen 3'-Enden komplementär sind, besitzt, gefolgt von der Synthese des zweiten Stranges der stochastischen DNA.

Nach einer anderen Ausführungsform werden die stochastischen Gene hergestellt durch Copolymerisation von Oligonukleotiden ohne kohäsive Enden, in einer Weise, um Fragmente stochastischer DNA zu bilden gefolgt von einer Ligierung dieser Fragmente an einen vorher linearisierten Expressionsvektor.

Der Expressionsvektor kann ein Plasmid sein, insbesondere ein bakterielles Plasmid. Hervorragende Ergebnisse wurden unter Verwendung des Plasmid pUC8 als Expressionsvektor erhalten.

Der Expressionsvektor kann auch eine virale DNA oder ein Hybrid eines Plasmids und einer viralen DNA sein.

Die Wirtszellen können prokaryontische Zellen wie z. B. HB 101 und C 600 oder eukaryontische Zellen sein.

Wenn das Verfahren gemäß der zweiten oben erwähnten Ausführungsform verwendet wird, ist es möglich, Oligonukleotide zu verwenden, die eine Gruppe palindromischer Octamerer bilden.

Insbesondere gute Ergebnisse werden erhalten unter Verwendung der folgenden Gruppen von palindromischer Octamerer:

5' GGAATTCC 3'

5' GGTCCACC 3'

5' CAAGCTTG 3'

5' CCATATCG 3'

5' CATCCATG 3'

Es ist auch möglich, Oligonukleotide zu verwenden, die eine Gruppe palindromischer Heptamerer bilden.

Sehr gute Ergebnisse werden erhalten unter Verwendung der folgenden Gruppe palindromischer Heptamerer:

5' XTCCCGA 3'

5' XCTCCAG 3'

5' RGCTACC 3'

worin X = A, G, C oder T und R = A oder T ist.

Nach einer Methode, um diese Verfahren zu verwenden, die besonders vorteilhaft ist, isoliert und reinigt man die transformierende DNA der Plasmide aus einer Kultur unabhängiger Klone der transformierten Wirtszellen, die nach den obigen Verfahren erhalten wurden, dann wird die gereinigte DNA durch mindestens ein Restriktionsenzym geschnitten, das spezifischen enzymatischen Schnittstellen entspricht, die in den palindromischen Octameren oder Heptameren vorhanden sind, aber in dem verwendeten Expressionsvektor fehlen; auf dieses Schneiden folgt Inaktivierung des Restriktionsenzyms, dann behandelt man das Ensemble der so erhaltenen linearisierten stochastischen DNA-Fragmente gleichzeitig mit T4-DNA-Ligase in einer solchen Weise, um ein neues Ensemble von DNA zu bilden, das neue stochastische Sequenzen enthält, wobei dieses neue Ensemble deshalb eine Zahl stochastischer Gene enthalten kann, die größer ist als die Zahl der Gene in dem anfänglichen Ensemble. Dieses neue Ensemble transformierender DNA verwendet man dann zur Transformation der Wirtszellen und Klonierung dieser Gene und schließlich wendet man Screening und/oder Selektion an und isoliert die neuen Klone der transformierten Wirtszellen und schließlich werden diese kultiviert, um die RNA zu produzieren.

Gemäß einer bevorzugten Ausführungsform ist das Kriterium zur Selektion der Klone transformierter Wirtszellen die Fähigkeit derselben bestimmte Peptide, Polypeptide oder Proteine zu exprimieren, die die Wirkung eines biologisch aktiven Moleküls simulieren oder modifizieren, z. B. eines Proteins. Das Screening und/oder die Selektion der Klone transformierter Wirtszellen, die ein solches Peptid, Polypeptid oder Protein produzieren, wird durchgeführt, indem man Antikörper gegen das aktive Molekül herstellt, dann diese Antikörper nach ihrer Reinigung verwendet, um die Klone, die dieses Peptid, Polypeptid oder Protein enthalten, zu identifizieren. Nach Kultivieren der so identifizierten Klone, Abtrennen und Reinigen der durch diese Klone produzierten Peptide, Polypeptide oder Proteine kann man das Peptid, Polypeptid oder Protein einer in vitro-Untersuchung unterwerfen, um sicherzustellen, daß es die Fähigkeit besitzt, die Wirkungen des Moleküls zu simulieren oder zu modifizieren.

Gemäß einer Abänderung des Verfahrens identifiziert und isoliert man die Klone transformierter Wirtszellen, die Peptide oder Polypeptide mit den gewünschten Eigenschaften produzieren, durch Affinitätschromatographie gegen Antikörper, die einem Protein, das durch den natürlichen Teil des DNA-Hybrids exprimiert ist, entsprechen.

Im Fall, bei dem der natürliche Teil der Hybrid-DNA ein Gen enthält, das β -Galactosidase exprimiert, kann man z. B. die Klone der transformierten Wirtszellen vorteilhafterweise durch Affinitätschromatographie gegen Anti- β -Galactosidase-Antikörper identifizieren und isolieren.

Gemäß einer vorteilhaften Ausführungsform des erfindungsgemäßen Verfahrens bestehen die Wirtszellen aus Bakterien wie z. B. *Escherichia coli*, dessen Genom weder das natürliche Gen, das β -Galactosidase exprimiert, noch das EBG-Gen enthält, d. h. Z^- , EBG $^-$ *E. coli*. Die transformierten Zellen werden in dem Medium in Gegenwart von X-gal und dem Indikator IPTG kultiviert, und Zellen, die für β -Galactosidase-Funktionen positiv sind, werden bestimmt; danach wird für eine Großkultur die transformierende DNA in ein geeignetes Klon von Wirtszellen transplantiert, um mindestens ein Peptid oder Polypeptid zu produzieren.

Die als Kriterium zur Selektion der transformierten Wirtszellen dienende Eigenschaft kann auch die Fähigkeit der durch Kultivierung dieser Klone produzierten Polypeptide oder Proteine sein, an eine bestimmte Verbindung zu binden.

Diese Verbindung kann vorteilhafterweise insbesondere ausgewählt sein unter Peptiden, Polypeptiden und Proteinen, insbesondere unter Proteinen, die die Transkriptionsaktivität der DNA regulieren.

Gegenstand der Erfindung ist ein Verfahren zur Herstellung von RNA, gekennzeichnet durch gleichzeitige Herstellung von Genen im gleichen Medium, die mindestens teilweise aus stochastischen synthetischen Polynukleotiden aufgebaut sind, daß die so erhaltenen Gene in Wirtszellen eingebracht werden, um ein Ensemble transformierter Wirtszellen herzustellen, daß die so hergestellten Wirtszellen gleichzeitig kultiviert werden und daß Screening und/oder Selektion dieses Ensembles in einer Weise durchgeführt wird, um die Wirtszellen zu identifizieren, die stochastische Sequenzen von RNA mit mindestens einer gewünschten Eigenschaft besitzen und daß die RNA aus den so identifizierten Wirtszellen isoliert wird.

Die genannte Eigenschaft kann die Fähigkeit sein, eine bestimmte Verbindung zu binden, die z. B. ein Peptid oder Polypeptid oder Protein sein könnte, oder auch die Fähigkeit, eine bestimmte chemische Reaktion zu katalysieren, oder die Fähigkeit, eine Transfer-RNA zu sein.

Das erfindungsgemäße Verfahren sowie einige seiner Anwendungen wird nun, unter Bezugnahme auf nicht einschränkende Ausführungsformen näher beschrieben.

Zunächst werden insbesondere brauchbare Verfahren beschrieben, um die Synthese von stochastischen Genen durchzuführen, und die Einführung dieser Gene in Bakterien, um Klone transformierter Bakterien zu bilden.

1) Direkte Synthese eines Expressionsvektors

a) Linearisierung des Vektors

30 μ g, das sind ca. 10^{13} Moleküle des pUC8 Expressionsvektors, werden durch Inkubation während 2 Stun-

den bei 37°C mit 100 Einheiten PstI-Restriktionsenzym in einem Volumen von 300 l des geeigneten Standardpuffers linearisiert. Der linearisierte Vektor wird mit Phenol-Chloroform behandelt, dann in Ethanol ausgefällt, in einem Volumen von 30 l aufgenommen und auf ein 0,8% Agarose-Gel in Standard TEB-Puffer aufgebracht. Nach Wanderung in einem Feld von 3V/cm während 3 Stunden wird der linearisierte Vektor elektroverdünn, in Ethanol ausgefällt und in 30 l Wasser aufgenommen.

b) Stochastische Synthese unter Verwendung des Enzyms Terminal Transferase (TdT)

30 μ g des linearisierten Vektors werden 2 Stunden lang bei 37°C mit 30 Einheiten TdT in 300 l des geeigneten Puffers in Gegenwart von 1 mM dGTP, 1 mM dCTP, 0,3 mM dTTP und 1 mM dATP reagieren gelassen. Die niedrigere Konzentration von dTTP wird gewählt, um die Häufigkeit von "Stop"-Codons in der entsprechenden Messenger RNA zu verringern. Ein ähnliches Resultat, obgleich etwas ungünstiger, kann erhalten werden durch Verwendung einer niedrigeren Konzentration von dATP als die der anderen Desoxynukleotid-Triphosphate. Der Verlauf der Polymerisation an dem 3'-Ende der PstI-Stellen wird durch Analyse aliquoter Anteile an einem Gel verfolgt, die während des Verlaufes der Reaktion entnommen werden.

Wenn die Reaktion einen mittleren Wert von 300 Nukleotiden pro 3'-Ende erreicht oder durchläuft, wird sie abgebrochen und die freien Nukleotide werden von dem Polymeren durch differenzielle Ausfällung oder durch Leiten über eine Säule, die ein Molekularsieb wie z. B. Biogel P60 enthält, getrennt. Nach Konzentrierung durch Ausfällung in Ethanol werden die Polymeren einer weiteren Polymerisation mit TdT unterworfen, zuerst in Gegenwart von dATP, dann in Gegenwart von dTTP. Diese letzten zwei Reaktionen werden durch eine Filtration an einem Gel getrennt und werden in kurzen Intervallen (30 Sekunden bis 3 Minuten) durchgeführt, um aufeinanderfolgend 10 bis 30 A, gefolgt von 10 bis 30 T den 3'-Enden der Polymeren zuzufügen.

c) Synthese des zweiten Stranges von stochastischer DNA

Jedes Molekül des Vektors besitzt an dem Ende des vorhergehenden Durchlaufes zwei stochastische Sequenzen, dessen 3'-Enden komplementär sind. Die Mischung der Polymeren wird deshalb unter Bedingungen inkubiert, die die Hybridisierung der komplementären Enden begünstigen (150 mM NaCl, 10 mM Tris-HCl, pH 7,6, 1 mM EDTA bei 65°C während 10 Minuten, gefolgt von einer Erniedrigung der Temperatur auf 22°C mit einer Geschwindigkeit von 3 bis 4°C pro Stunde). Die hybridisierten Polymeren werden dann mit 60 Einheiten des großen Fragmentes (Klenow) von Polymerase 1 in Gegenwart der vier Nukleotid-Triphosphate (200 mM) bei 4°C während 2 Stunden umgesetzt. Diese Stufe bewirkt die Synthese des zweiten Stranges aus den 3'-Enden der Hybrid-Polymeren. Die Moleküle, die aus dieser direkten Synthese, ausgehend vom linearisierten Vektor, resultieren, werden danach verwendet, um kompetente Zellen zu transformieren.

d) Transformierung kompetenter Klone

100 bis 200 ml von kompetentem HB101 von C600

werden bei der Konzentration von 10^{10} Zellen/ml mit der stochastischen DNA-Präparation (von oben) in Gegenwart von 6 mM CaCl_2 , 6 mM Tris-HCl, pH 8, 6 mM MgCl_2 30 Minuten bei 0°C inkubiert. Die Mischung wird einem Temperaturschock von 3 Minuten bei 37°C unterworfen, gefolgt von der Zugabe von 400 bis 800 ml NZY Kulturmedium ohne Antibiotika.

Die transformierte Kultur wird bei 37°C während 16 Minuten inkubiert, dann auf 10 l durch Zugabe von NZY-Medium, das 40 $\mu\text{g/ml}$ Ampicillin enthält, verdünnt. Nach 3 bis 5 Stunden Inkubation bei 37°C wird die verstärkte Kultur zentrifugiert und die Pellets der transformierten Zellen lyophilisiert und bei -70°C aufbewahrt. Eine solche Kultur enthält 3×10^7 bis 10^8 unabhängige Transformanten wobei jeder ein einziges stochastisches Gen enthält, eingefügt in den Expressionsvektor.

II) Synthese stochastischer Gene, ausgehend von Oligonukleotiden ohne kohäsive Enden

Dieses Verfahren basiert auf der Tatsache, daß die Polymerisation von vernünftig gewählten palindromischen Oligonukleotiden den Aufbau von stochastischen Zellen ermöglicht, die kein "Stop"-Codon in einem der 6 möglichen Leserahmen besitzen, während sie gleichzeitig eine ausgeglichene Repräsentation von Triplets, die alle Aminosäuren spezifizieren, sicherstellen. Um eine Wiederholung von Sequenzmotiven in den entstehenden Proteinen zu vermeiden, können die Oligonukleotide weiters eine Zahl von Basen enthalten, die kein vielfaches von 3 ist. Das folgende Beispiel beschreibt die Verwendung einer der möglichen Kombinationen, die diese Kriterien erfüllen:

a) Wahl einer Gruppe von Oktameren

Die Gruppe der folgenden Oligonukleotide:

```

5'   CGAATTCC   3'
5'   GGTCCACC   3'
5'   CAAGCTTG   3'
5'   CCATATCG   3'
5'   CATCGATG   3'

```

setzt sich aus 5 Palindromen (also selbstkomplementären Sequenzen) zusammen, bei denen es leicht ist sicherzustellen, daß ihre stochastische Polymerisation nicht irgendwelche "Stop"-Codons hervorruft, und spezifiziert alle Aminosäuren.

Man kann selbstverständlich andere Gruppen palindromischer Octamerer verwenden, die nicht irgendwelche "Stop"-Codons hervorrufen und alle Aminosäuren spezifizieren, die in Polypeptiden gefunden werden. Es ist selbstverständlich auch möglich, nicht palindromische Gruppen von Octameren zu verwenden unter der Bedingung, daß ihre Komplemente, die doppelsträngige DNA bilden, auch verwendet werden.

b) Aufbau eines stochastischen Gens aus einer Gruppe von Octameren

Eine Mischung, die 5 μg jedes der oben angegebenen Oligonukleotide enthält (vorher an der 5'-Stelle durch

ein Standardverfahren phosphoryliert) wird in einem Volumen von 100 μl , das 1 mM ATP, 10% Polyethylenglykol und 100 Einheiten T4 DNA-Ligase in dem geeigneten Puffer enthält, bei 13°C während 6 Stunden umgesetzt. Diese Stufe bewirkt die stochastische Polymerisation der Oligomeren im doppelsträngigen Zustand ohne kohäsive Enden. Die entstehenden Polymeren werden durch Laufenlassen über ein Molekularsieb (Biogel P60) isoliert, wobei diejenigen mit 20 bis 100 Oligomeren erhalten werden. Nach Konzentrierung wird diese Fraktion wieder einer Katalyse oder Polymerisation durch T4 DNA-Ligase unter den oben beschriebenen Bedingungen unterworfen. Danach werden, wie oben beschrieben, diejenigen Polymere isoliert, die aus mindestens 100 Oligomeren aufgebaut sind.

c) Herstellung des Wirtsplasmids

Der pUC8-Expressionsvektor wird durch SmaI-Enzym in dem geeigneten Puffer, wie oben beschrieben, linearisiert. Der durch SmaI linearisierte Vektor besitzt keine kohäsiven Enden. Der so linearisierte Vektor wird mit alkalischer Kälbermagenphosphatase (CIP) bei einem Gehalt von einer Einheit pro mg des Vektors in dem geeigneten Puffer bei 37°C 30 Minuten lang behandelt. Das CIP-Enzym wird danach mittels zweier aufeinanderfolgender Extraktionen mit Phenol-Chloroform inaktiviert. Der linearisierte und dephosphorylierte Vektor wird in Ethanol ausgefällt, dann mit 1 mg/ml in Wasser wieder gelöst.

d) Ligierung von stochastischen Genen an den Vektor

Äquimolare Mengen des Vektors und Polymers werden gemischt und in Gegenwart von 1000 Einheiten T4 DNA-Ligase, 1 mM ATP, 10% Polyethylenglykol in dem geeigneten Puffer 12 Stunden lang bei 13°C inkubiert. Diese Stufe ligiert die stochastischen Polymeren in den Expressionsvektor und bildet doppelsträngige, kreisförmige Moleküle, die deshalb zur Transformation geeignet sind.

Transformation kompetenter Klone

Die Transformation kompetenter Klone wird in der vorstehend beschriebenen Weise durchgeführt.

III) Aufbau stochastischer Gene, ausgehend von einer Gruppe von Heptameren

Dieses Verfahren unterscheidet sich von dem gerade beschriebenen dadurch, daß man palindrome Heptamer verwendet, die variable kohäsive Enden besitzen, anstelle der stochastischen Sequenzen, die eine geringere Anzahl identischer Motive enthalten.

a) Wahl einer Gruppe von Heptameren

Es ist, als ein Beispiel, möglich, die folgenden drei palindromischen Heptameren zu verwenden:

```

5'   XTCCGCA   3'
5'   XCTGCAG   3'
5'   RCGTACC   3'

```

worin X = A, G, C, oder T und R = A oder T, und worin

die Polymerisation nicht irgendwelche "Stop"-Codons bilden kann und Triplets bildet, die alle Aminosäuren spezifizieren. Es ist selbstverständlich möglich, andere Gruppen von Heptameren zu verwenden, die diese gleichen Bedingungen erfüllen.

b) Polymerisation einer Gruppe von Heptameren

Diese Polymerisation wird in genau dergleichen Weise durchgeführt, wie sie oben für Octamere beschrieben ist.

c) Eliminierung kohäsiver Enden

Die so erhaltenen Polymeren haben an ihren zwei 5'-Enden eine ungepaarte Base. Es ist deshalb notwendig, die komplementäre Base an den entsprechenden 3'-Enden anzufügen. Dies wird wie folgt ausgeführt: 10 µg der doppelsträngigen Polymeren werden mit 10 Einheiten des Klenow-Enzyms in Gegenwart der 4-Deoxynukleotid-phosphate (200 mM) in einem Volumen von 100 µl bei 4°C während 60 Minuten umgesetzt. Das Enzym wird durch Phenol-Chloroform-Extraktion inaktiviert und die Polymeren werden von den verbliebenen freien Nukleotiden durch differentielle Ausfällung gereinigt. Die Polymeren werden dann an das Wirtsplasmid (vorher linearisiert und dephosphoryliert) gemäß dem oben beschriebenen Verfahren ligiert.

Es ist festzustellen, daß die zwei letztbeschriebenen Verfahren palindromische Octamere oder Heptamere verbinden, die spezifische Stellen von gewissen Restriktionsenzymen bilden. Diese Stellen fehlen zum größten Teil im pUC8-Expressionsvektor. Es ist deshalb möglich, die Komplexität einer anfänglichen Präparation von stochastischen Genen beträchtlich zu vermehren, wenn man nach dem folgenden Weg vorgeht: Die Plasmid-DNA, die sich von der Kultur von 10^7 unabhängigen Transformanten, die nach einer der beiden letzten oben beschriebenen Verfahren erhalten wurde, ableitet, wird isoliert. Nachdem diese DNA gereinigt ist, wird sie teilweise von ClaI-Restriktionsenzym (Verfahren II) oder von Pst I-Restriktionsenzym (Verfahren III) verdaut. Nach Inaktivierung des Enzyms wird die teilweise verdaut DNA mit T4-DNA-Ligase behandelt, die die Wirkung hat, eine sehr große Zahl neuer Sequenzen hervorzurufen, während die fundamentalen Eigenschaften der anfänglichen Sequenzen bewahrt bleiben. Dieses neue Ensemble stochastischer Sequenzen kann dann verwendet werden, um kompetente Zellen zu transformieren. Zusätzlich können die nach dem Verfahren II und III klonierten stochastischen Gene in intakter Form aus dem pUC8-Expressionsvektor unter Verwendung von Restriktionsstellen herausgeschnitten werden, die zu dem klonierenden Vektor gehören und in den stochastischen DNA-Sequenzen nicht vorhanden sind.

Die Rekombinierung innerhalb der durch die zwei gerade beschriebenen Verfahren gebildeten stochastischen Gene, die aus der internen Homologie aufgrund der wiederkehrenden molekularen Motive resultieren, ist eine wichtige zusätzliche Methode, um in vivo Mutagenese der kodierenden Sequenzen zu erreichen. Dies ergibt eine Vermehrung der Zahl neuer Gene, die geprüft werden können.

Schließlich ist es für alle Verfahren zur Schaffung neuer synthetischer Gene möglich, eine Zahl allgemein üblicher Verfahrensweisen zu verwenden, um Gene in vivo oder in vitro zu modifizieren, wie z. B. eine Änderung des Leserahmens, Inversion von Sequenzen im

Hinblick auf ihren Promotor, Punktmutationen oder Verwendung von Wirtszellen, die eine oder mehrere Suppressor-tRNAs exprimieren.

Unter Berücksichtigung der obigen Beschreibung ist es klar, daß es möglich ist, in vitro eine extrem große Zahl (z. B. mehr als eine Billion) verschiedener Gene durch enzymatische Polymerisation von Nukleotiden oder von Oligonukleotiden zu konstruieren. Diese Polymerisation wird in einer stochastischen Weise durchgeführt, wie sie bestimmt wird durch die jeweiligen Konzentrationen der Nukleotide oder Oligonukleotide, die in der Reaktionsmischung vorhanden sind.

Wie vorstehend gezeigt, können zwei Verfahren verwendet werden, um solche Gene (oder kodierende Sequenzen) zu klonieren: die Polymerisation kann direkt an einem klonierenden Expressionsvektor durchgeführt werden, der vorher linearisiert wurde; oder es ist möglich, auf die Polymerisation dann die Ligierung der Polymeren auf dem Expressionsvektor folgen zu lassen.

In den zwei Fällen ist die nächste Stufe die Transformation oder Transvektion kompetenter Bakterienzellen (oder Zellen in Kultur). Diese Stufe konstituiert die Klonierung der stochastischen Gene in lebenden Zellen, wo sie unbegrenzt vermehrt und exprimiert werden.

Es ist klar, daß zusätzlich zu den vorstehend beschriebenen Verfahren es möglich ist, alle anderen Verfahren zu verwenden, die geeignet sind zur Synthese stochastischer Sequenzen. Insbesondere ist es möglich, die Polymerisation von einzelsträngigen Oligomeren der DNA oder RNA, erhalten durch chemische Synthese, durch biochemische Mittel durchzuführen, dann diese Segmente von DNA oder RNA nach bewährten Verfahren zu behandeln, um doppelsträngige DNA (cdNA) auszubilden, um solche Gene zu klonieren.

Screening oder Selektion von Klonen transformierter Wirtszellen

Die weitere Stufe des erfindungsgemäßen Verfahrens besteht in der Prüfung der transformierten oder transfektierten Zellen durch Selektion oder Screening, um eine oder mehrere Zellen zu isolieren, deren transformierende oder transfektierende DNA zur Synthese eines Transkriptionsproduktes (RNA) oder Translationsproduktes (Protein) führt, das die gewünschte Eigenschaft besitzt. Diese Eigenschaften können z. B. enzymatischer, funktioneller oder struktureller Natur sein.

Einer der wichtigsten Aspekte des erfindungsgemäßen Verfahrens ist es, daß es das gleichzeitige Screening oder Selektion eines verwertbaren Produktes (RNA) oder Protein) und des Gens, welches dieses Produkt produziert, ermöglicht.

Wir werden nun, als nicht einschränkende Beispiele, bevorzugte Verfahren zum Screening oder zur Selektion von Klonen transformierter Zellen beschreiben, die solcher Art sind, daß die neuen Proteine vom Standpunkt industrieller oder medizinischer Anwendungen von Interesse sind.

Eines dieser Verfahren beruht auf der Idee der Herstellung oder des Erhalts polyklonaler oder monoklonaler Antikörper durch bewährte Techniken, die gegen ein Protein oder einen anderen Typ eines Moleküls von biochemischem oder medizinischem Interesse gerichtet sind, wobei dieses Molekül immunisierend ist oder gemacht wurde und nachfolgende Verwendung dieser Antikörper als Proben, um unter der sehr großen Zahl von Klonen, die durch stochastische Gene transformiert wurden, diese zu identifizieren, dessen Protein mit die-

sen Antikörpern reagiert. Diese Reaktion ist ein Ergebnis einer strukturellen Homologie, die zwischen dem durch die stochastischen Gene synthetisierten Polypeptid und dem anfänglichen Molekül besteht. Es ist auf diese Weise möglich, zahlreiche neue Proteine zu isolieren, die als epitope oder antigene Determinanten auf dem anfänglichen Molekül wirken. Solche neuen Proteine neigen dazu, die Wirkung des anfänglichen Moleküls zu simulieren, stimulieren, modulieren oder zu blockieren. Es ist verständlich, daß dieses Mittel von Selektion oder Screening selbst sehr viele pharmakologische und biochemische Anwendungen haben kann. Nachfolgend beschreiben wir als nicht einschränkendes Beispiel diese erste Verfahrensart in einem konkreten Fall:

EGF (epidermal growth factor) ist ein kleines Protein, das im Blut vorhanden ist, dessen Rolle es ist, das Wachstum von Epithelzellen zu stimulieren. Diese Wirkung wird durch die Wechselwirkung von EGF mit einem spezifischen Rezeptor erhalten, der sich in der Membran der Epithelzellen befindet.

Gegen EGF gerichtete Antikörper werden hergestellt, indem man Tiere mit an KLH (keyhole limpet hemocyanin) gekoppeltes EGF in Tiere injiziert, um die Immunität des EGF zu erhöhen. Die Anti-EGF-Antikörper der immunisierten Tiere werden gereinigt, z. B. durch Leiten über eine Affinitätsäule, worin der Ligand EGF oder ein synthetisches Peptid, das einem Fragment von EGF entspricht, ist. Diese gereinigten Anti-EGF-Antikörper werden an einem festen Träger als Proben verwendet, um eine große Zahl von durch Chloroform gelysten Bakterienklone zu screenen. Die Anti-EGF-Antikörper binden solche stochastischen Peptide oder Proteine, deren Epitope denen des anfänglichen Antigens gleichen. Die solche Peptide oder Proteine enthaltenden Klone werden durch Autoradiographie nach Inkubation des festen Trägers mit radioaktivem Protein A, oder nach Inkubation mit einem radioaktiven Anti-Antikörper-Antikörper sichtbar gemacht.

Diese Stufen identifizieren solche Klone, von denen jedes ein Protein (und sein Gen) enthält, das mit dem screenenden Antikörper reagiert. Es ist möglich, unter einer sehr großen Zahl von Bakterienzellen-Kolonien oder viralen Plaques (z. B. in der Größenordnung von einer Million) zu screenen und es ist möglich, extrem kleine Mengen des Proteinproduktes, in der Größenordnung von 1 ng, festzustellen. Danach werden die identifizierten Klone kultiviert und die so bestimmten Proteine werden auf konventionellen Wegen gereinigt. Diese Proteine werden in vitro in Epithelzellen-Kulturen getestet um zu bestimmen, ob sie die Wirkungen von EGF auf diese Kulturen inhibieren, simulieren oder modulieren. Unter den so erhaltenen Proteinen können einige für die chemotherapeutische Behandlung von Hautkrebs verwendet werden. Die Wirkungen der so erhaltenen Proteine können durch Mutation der für die Proteine kodierenden DNA verbessert werden auf Wegen, die denen der oben beschriebenen analog sind. Eine Variante dieses Verfahrens besteht in der Reinigung dieser stochastischen Peptide, Polypeptide oder Proteine, die als Impfstoffe verwendet werden können, oder allgemeiner dazu, eine Immunität gegen ein pathogenes Mittel zu übertragen oder andere Wirkungen auf das immunologische System auszuüben, z. B. um im Hinblick auf ein bestimmtes Antigen eine Toleranz zu schaffen oder die Überempfindlichkeit zu verringern, insbesondere aufgrund der Bindung dieser Peptide, Polypeptide oder Proteine mit den gegen dieses Antigen gerichteten Antikörpern. Es ist klar, daß es möglich

ist, solche Peptide, Polypeptide oder Proteine in vitro sowie in vivo zu verwenden.

Genauer ausgedrückt hat in dem Ensemble neuer Proteine, die mit den Antikörpern gegen ein bestimmtes Antigen X reagieren, jedes mindestens ein Epitop zusammen mit X, so daß das Ensemble ein Ensemble von Epitopen zusammen mit X besitzt. Dies erlaubt die Verwendung des Ensembles oder Sub-Ensembles als Impfstoff, um Immunität gegen X zu übertragen. Es ist z. B. einfach, ein oder mehrere der Capsid-Proteine des Hepatitis B Virus zu reinigen. Diese Proteine können dann in ein Tier injiziert werden, z. B. ein Kaninchen, und die dem anfänglichen Antigen entsprechenden Antikörper können durch Affinitätsäulen-Reinigung gewonnen werden. Diese Antikörper können, wie oben beschrieben, verwendet werden, um Klone zu identifizieren, die mindestens ein Protein produzieren, das ein Epitop besitzt, das mindestens einem der Epitope des anfänglichen Antigens ähnelt. Nach Reinigung werden diese Proteine als Antigene verwendet (entweder allein oder in Kombination), mit dem Ziel, Schutz gegen Hepatitis B zu übertragen. Die endgültige Herstellung des Impfstoffes erfordert keinen weiteren Zugang zu dem anfänglichen pathogenen Mittel.

Nachfolgend sind die Prinzipien und Verfahren zur Selektion von Peptiden oder Polypeptiden und der entsprechenden Gene beschrieben gemäß einem zweiten Verfahren des Screenings und der Selektion, das auf der Bestimmung der Fähigkeit dieser Peptide oder Polypeptide beruht, eine spezifische Reaktion zu katalysieren.

Als ein konkretes und nicht einschränkendes Beispiel wird das Screening oder die Selektion in dem besonderen Fall eines Proteins beschrieben, das fähig ist, die Spaltung von Lactose zu katalysieren, eine Funktion, die normalerweise von dem Enzym β -Galactosidase (β -gal) erfüllt wird.

Wie vorstehend beschrieben, besteht die erste Stufe des Verfahrens in der Schaffung eines sehr großen Ensembles von Expressionsvektoren von denen jeder ein bestimmtes neues Protein exprimiert. Im konkreten Fall kann man z. B. den pUC8-Expressionsvektor wählen, mit Klonieren stochastischer Sequenzen von DNA in der PstI-Restriktionsstelle. Die so erhaltenen Plasmide werden dann in ein Klon von *E. coli* eingebracht, von dessen Genom das natürliche Gen für β -Galactosidase, Z und ein zweites Gen EBG, das zum ersten in keiner Beziehung steht aber dazu fähig ist, in Richtung β -gal-Funktion zu mutieren, beide durch bekannte genetische Methoden entfernt wurden. Solche Wirtszellen (Z^- , EBG $^-$) sind für sich nicht fähig, die Lactose-Hydrolyse zu katalysieren und als Konsequenz Lactose als Kohlenstoffquelle für das Wachstum zu verwenden. Dies erlaubt die Verwendung solcher Wirtsklone für das Screening oder die Selektion für β -gal-Funktion.

Eine zweckmäßige biologische Prüfung zur Analyse transformierter *E. coli* Klone auf solche, die neue Gene, die eine β -gal-Funktion exprimieren, besteht in der Kultur von wie beschrieben transformierten Bakterien in Petrischalen, die X-gal in dem Medium enthalten. In diesem Fall werden alle Bakterienkolonien, die eine β -gal-Funktion exprimieren, als blaue Kolonien sichtbar gemacht. Unter Verwendung einer solchen biologischen Prüfung ist es möglich, sogar schwache katalytische Aktivität festzustellen. Die spezifische Aktivität von charakteristischen Enzymen liegt im Bereich von 10 bis 10 000 Produktmolekülen pro Sekunde.

Unter der Annahme, daß ein durch ein stochastisches Gen synthetisiertes Protein eine schwache spezifische

Aktivität besitzt in der Größenordnung von einem Molekül pro 100 Sekunden, bleibt es möglich, eine solche katalytische Aktivität zu bestimmen. In einer Petrischale, die im Medium X-gal enthält und in Gegenwart des nicht metabolisierbaren Inducer IPTG (Isopropyl-D-thiogalactosid) erfordert die Sichtbarmachung einer blauen Region die Spaltung von ca. 10^{10} bis 10^{11} Molekülen an X-gal pro mm^2 . Eine Bakterienkolonie, die ein schwaches Enzym exprimiert und eine Oberfläche von 1 mm^2 einnimmt, hat ca. 10^7 bis 10^8 Zellen. Wenn jede Zelle nur eine Kopie des schwachen Enzyms besitzt, würde jede Zelle zur Katalyse der Spaltung zwischen 10 000 und 100 von X-gal benötigen, um bestimmt zu werden, was zwischen 2,7 und 270 Stunden erfordern würde. Da es unter selektiven Bedingungen möglich ist, die Zahl der Kopien jedes Plasmids pro Zelle auf 5 bis 20 Kopien pro Zelle zu verstärken oder sogar auf 100 oder 1000 und weil bis zu 10% des Proteins der Zelle durch das neue Gen spezifiziert werden kann, ist die zur Bestimmung einer blauen Kolonie im Fall von 100 Enzymmolekülen schwacher Aktivität pro Zelle benötigte Zeit in der Größenordnung von 0,27 bis 2,7 Stunden.

Als Konsequenz dieser Tatsachen ist das Screening einer sehr großen Zahl unabhängiger Bakterienkolonien, von denen jede ein verschiedenes neues Gen exprimiert und die Verwendung der Fähigkeit, eine β -gal-Funktion zu exprimieren, als Selektionskriterium voll möglich. Es ist möglich, das Screening von ca. 2000 Kolonien in einer Petrischale von 10 cm Durchmesser durchzuführen. So können ca. 20 Millionen Kolonien auf einem Blatt von X-gal-Agar von 1 m^2 gescreent werden.

Es ist festzustellen, daß Bakterienkolonien, die auf X-gal-Petrischalen blau erscheinen, falsche Positive sein könnten, aufgrund einer Mutation in dem Bakteriengenom, die die Fähigkeit, Lactose zu metabolisieren, darauf überträgt oder aus anderen Gründen als denen, die aus einer katalytischen Wirkung des neuen, durch die Zellen der Kolonie exprimierten Proteins resultieren. Solche falschen Positive können direkt eliminiert werden durch Reinigung der DNA des Expressionsvektors von der positiven Kolonie und Retransformation von Z^- , EBG $^-$, E. coli-Wirtszellen. Wenn die β -gal-Aktivität auf dem durch das neue Gen in dem Expressionsvektor kodierte Protein beruht, werden alle die durch diesen Vektor transformierten Zellen β -gal-Funktion zeigen. Im Gegensatz dazu ist es eine seltene Erscheinung und unabhängig von der Transformation, wenn die anfängliche blaue Kolonie auf einer Mutation im Genom der Wirtszelle beruht und deshalb wird die Zahl der Zellen des neuen Klon von transformiertem E. coli, die zur Expressierung der β -gal-Funktion fähig sind, klein oder 0 sein.

Die Leistung der gleichzeitigen Massenreinigung aller Expressionsvektoren all den positiven Klonen (blau), gefolgt von Retransformation naiver Bakterien sollte hervorgehoben werden. Unter der Annahme, daß es das Ziel ist, ein Screening durchzuführen, um Proteine zu selektieren, die eine katalytische Funktion besitzen und daß die Wahrscheinlichkeit, daß ein neues Peptid oder Polypeptid diese Funktion mindestens wöchentlich ausführt, ist 10^{-6} , während die Wahrscheinlichkeit, daß ein Klon des bakteriellen E. coli-Wirts einer Mutation unterliegt, die ihn dazu befähigt, die gleiche Funktion auszuüben, 10^{-5} ist, dann kann errechnet werden, daß unter 20 Millionen transformierten Bakterien, die gescreent werden, 20 positive Klone den neuen Genen in den Expressionsvektoren, welche jedes tragen, zuzuschreiben sind, während 200 positive Klone das Ergebnis einer

Genom-Mutation sein werden. Die Massenreinigung der Expressionsvektoren von den insgesamt 220 positiven bakteriellen Klonen, gefolgt von der Retransformation naiver Bakterien mit der Mischung dieser Expressionsvektoren wird eine große Zahl positiver Klone produzieren, die aus allen diesen Bakterien bestehen, die mit den 20 Expressionsvektoren transformiert sind, die für die neuen Proteine kodieren, die die gewünschte Funktion besitzen und eine sehr kleine Zahl von bakteriellen Klonen, die aus genomischen Mutationen resultieren und die 200 Expressionsvektoren enthalten, die nicht von Interesse sind. Eine kleine Zahl von Reinigungszyklen von Expressionsvektoren von positiven bakteriellen Kolonien, gefolgt von einer solchen Retransformation, erlaubt die wirklich positive Bestimmung sehr seltener Expressionsvektoren für eine gewünschte katalytische Aktivität, trotz einer hohen Hintergrundrate von Mutationen in den Wirtszellen für die gleiche Funktion.

Es ist möglich, dieses Screeningverfahren für irgendeine enzymatische Funktion anzuwenden, für die eine geeignete biologische Prüfung existiert. Für solche Screenings ist es nicht notwendig, daß die enzymatische Funktion, die gesucht wird, für die Wirtszelle brauchbar ist. Es ist möglich, die Screenings nicht nur für eine enzymatische Funktion auszuführen, sondern für irgendeine andere gewünschte Eigenschaft, für die es möglich ist, eine geeignete biologische Prüfung zu schaffen. Es ist deshalb selbst in dem einfachen Fall einer β -gal-Funktion, die an einer X-gal-Petriplatte sichtbar gemacht wird möglich, ein Screening in der Größenordnung von 100 Millionen oder sogar einer Billion neuer Gene für eine katalytische Aktivität oder irgendeine andere gewünschte Eigenschaft durchzuführen.

Selektion transformierter Wirtszellen

Auf der anderen Seite ist es möglich, die Selektionstechniken für irgendeine Eigenschaft, katalytischer oder anderer Natur zu verwenden, worin die Gegenwart oder Abwesenheit der Eigenschaft für das Überleben der Wirtszellen, die die Expressionsvektoren enthalten, die für die neuen Gene kodieren, wesentlich ist oder auch verwendet werden kann, um diese Viren zu selektieren, die das gewünschte neue Gen kodieren und exprimieren. Als nicht eingeschränktes, aber konkretes Beispiel soll die Selektion der β -Galactosidase-Funktion beschrieben werden. Ein geeignetes Klon von Z^- EBG $^-$ E. coli ist nicht fähig, an Lactose als einzige Kohlenstoffquelle zu wachsen. Es ist deshalb nach Durchführung der ersten oben beschriebenen Stufe möglich, eine sehr große Zahl von Wirtszellen, die durch die Expressionsvektoren transformiert sind, die für die neuen Gene kodieren, unter selektiven Bedingungen zu kultivieren, entweder durch progressive Verringerung anderer Kohlenstoffquellen oder alleiniger Verwendung von Lactose vom Start an. Während der Durchführung einer solchen Selektion erlaubt die in vivo Mutagenese durch Rekombination oder durch explizite Gewinnung der Expressionsvektoren und Mutagenese ihrer neuen Gene in vitro durch verschiedene Mutagenen oder durch irgendeine andere übliche Technik, anpassungsfähige Verbesserungen in der Fähigkeit, die gewünschte katalytische Funktion zu erfüllen. Wenn zur gleichen Zeit Selektionstechniken und zweckmäßige Bioassay-Techniken existieren, wie im vorliegenden Fall, ist es möglich, die Selektionstechniken anfangs zu verwenden, um die

Repräsentation der Wirtsbakterien, die die β -gal-Funktion exprimieren, anzureichern und dann ein Screening auf einer Petriplatte an X-gal-Medium durchzuführen, um wirkungsvoll zu bestimmen, welches die positiven Zellen sind. In Abwesenheit zweckmäßiger Bioassays ist die Anwendung einer progressiv genaueren Selektion der einfachste Weg, um ein oder eine kleine Zahl bestimmter Wirtszellen zu reinigen, deren Expressionsvektoren für die Proteine kodieren, die die gewünschte Reaktion katalysieren.

Es ist möglich, diese Techniken zu verwenden, um neue Proteine zu finden, die eine große Vielzahl struktureller und funktioneller Charakteristika neben der Fähigkeit, eine spezifische Reaktion zu katalysieren, besitzen. Zum Beispiel ist es möglich, ein Screening oder Selektion für neue Proteine durchzuführen, die an cis-regulative Stellen an der DNA binden, und dabei die Expression einer der Wirtszellen-Funktionen blockieren oder die Transkription der DNA blockieren, die Transkription stimulieren usw.

Im Falle von *E. coli* exprimiert z. B. ein Klonmutant in dem Repressor des Lactose-Operons (*i*-) grundlegend die β -gal-Funktion aufgrund der Tatsache, daß der Lactose-Operator nicht reprimiert ist. Alle Zellen dieses Typs produzieren an Petriplatten, die X-gal-Medium enthalten, blaue Klone. Es ist möglich, solche Wirtszellen mit Expressionsvektoren zu transformieren, die neue Proteine synthetisieren und ein Screening an X-gal-Petriplatten durchzuführen, um diese Klone zu bestimmen, die nicht blau sind. Unter diesen repräsentieren einige den Fall, bei dem das neue Protein an den Lactose-Operator bindet und die Synthese von β -gal unterdrückt (reprimiert). Es ist dann möglich, solche Plasmide in Masse zu isolieren, zu retransformieren, solche Klone, die kein β -gal produzieren, zu isolieren und danach eine detaillierte Prüfung durchzuführen.

Wie vorstehend erwähnt, kann das Verfahren verwendet werden, um nicht nur verwertbare Proteine zu schaffen und dann zu isolieren, sondern auch RNA und DNA als Produkte, die selbst verwertbare Eigenschaften besitzen. Dies ergibt sich aus der Tatsache, daß auf der einen Seite das Verfahren darin besteht, stochastische Sequenzen von DNA zu schaffen, die direkt mit anderen zellulären oder biochemischen Bestandteilen in Wechselwirkung treten können und auf der anderen Seite diese in Expressionsvektoren klonierten Sequenzen in RNA transkribiert werden, die selbst für viele biochemische Wechselwirkungen fähig sind.

Schaffung und Selektion einer als solche verwertbaren RNA

Es wird eine sehr große Zahl stochastischer DNA-Sequenzen angenommen, die wie beschrieben hergestellt und in einem Expressionsvektor kloniert wurden. Es folgt, daß die aus diesen Sequenzen in den transformierten Wirtszellen transkribierte RNA als solche brauchbare Produkte sein können. Als nicht einschränkendes Beispiel ist es möglich, ein stochastisches Gen, das für eine Suppressor-Transfer-RNA (tRNA) kodiert, nach dem folgenden Verfahren zu selektieren: eine große Zahl ($\geq 10^8$) stochastischer Sequenzen wird in kompetente bakterielle Wirte transferiert, die eine "Nonsense"-Mutation in dem *arg* E. Gen tragen. Diese transformierten Bakterien werden im geringstmöglichen Medium ohne Arginin und mit dem selektiven Antibiotikum für das Plasmid (Ampicillin, wenn der Vektor pUC8 ist) auf eine Platte aufgeschichtet. Nur diejenigen transfor-

mierten Bakterien, die zur Synthese von Arginin befähigt wurden, sind fähig, zu wachsen. Dieser Phenotyp kann entweder von einer Rückmutation des Wirtsgenoms resultieren oder von der Gegenwart in der Zelle eines Suppressors. Es ist leicht, jede transformierte Kolonie zu testen, um festzustellen, ob der *arg*⁺-Phenotyp von der Gegenwart des stochastischen Gens in seinem Vektor abhängig ist oder nicht; es reicht aus, das Plasmid aus seiner Kolonie zu reinigen um festzustellen, daß es einen *arg*⁺-Phenotyp an allen dadurch transformierten *arg* E. Zellen trägt.

Patentansprüche

1. Verfahren zur Herstellung von RNA mit einer vorgegebenen Eigenschaft, welches folgende Stufen umfaßt:
in dem gleichen Milieu werden Gene gleichzeitig produziert;
die so produzierten Gene werden in Wirtszellen eingebracht, um ein Ensemble transformierter Wirtszellen zu bilden;
die so hergestellten unabhängigen Klone transformierter Wirtszellen werden gleichzeitig vermehrt;
es wird ein Screening und/oder eine Selektion an diesem Ensemble durchgeführt, um die Wirtszellen zu identifizieren, die die RNA mit der vorgegebenen Eigenschaft enthalten;
die RNA aus den Kulturen der identifizierten Wirtszellen wird isoliert,
dadurch gekennzeichnet, daß die Gene, die in die Wirtszellen eingebracht werden, mindestens teilweise aus synthetischen, stochastischen Polynukleotiden zusammengesetzt sind.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Eigenschaft die Fähigkeit ist, an eine bestimmte Verbindung zu binden.
3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Eigenschaft die Fähigkeit ist, eine bestimmte chemische Reaktion zu katalysieren.
4. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß es die Eigenschaft ist, eine Transfer-RNA zu sein.